



# Statistical Sciences Group Overview

Sallie Keller-McNulty, Ph. D.  
*sallie@lanl.gov*, 505-665-3957

***<http://www.stat.lanl.gov>***



# Los Alamos National Laboratory

## **Vision:**

We serve the nation by developing and applying the best science and technology to make the world a better and safer place.

## **Mission:**

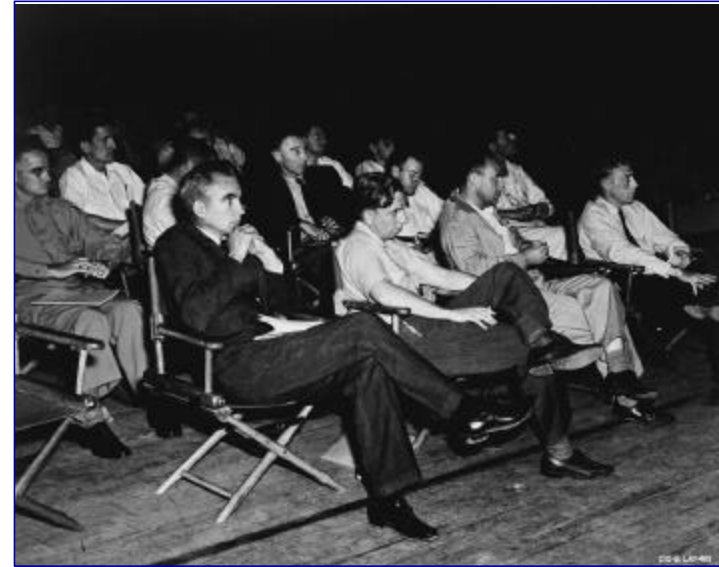
- Ensure the safety and reliability of the U.S. Nuclear deterrent.
- Reduce the global threat of weapons of mass destruction.
- Solve national problems in energy, environment, infrastructure and health security.

LANL is operated by the University of California for the U.S. Department of Energy





# Los Alamos 1945





*People*







# Los Alamos 2002





# Statistical Sciences Group

## **Mission:**

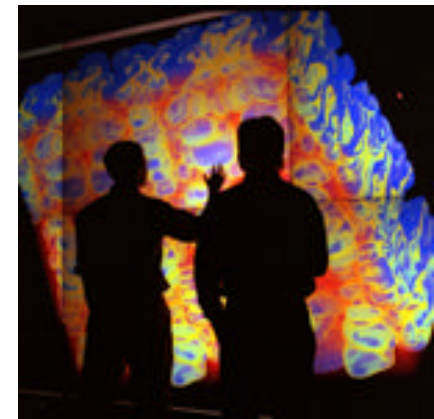
Bring statistical reasoning and rigor to multi-disciplinary scientific investigations through development, application, and communication of cutting-edge statistical sciences research.

## **Group Composition:**

- 17 Ph. D. Statisticians
- 3 Ph. D. Cultural Anthropology, Sociology, Rhetoric
- 4 MS/BS Statisticians
- 1 Ph. D., 1 BS Computer Scientists
- 2 Electronic Communications support staff
- Office support and system administration

## **Special Programs:**

- Post Docs
- GRAs
- Visiting Faculty
- Workshops and Conferences





# How do we work?

---

- Both basic and applied statistical research problems, broadly defined.
- Work on funded projects, some led within the group and some by other groups.
- Funding is strong, thus allowing us to pick and choose which projects to work on as well as directions for the group.
  - ◆ Current Funding Sources: DOE, DOD, NIH, NCES, NSF, Industrial Partners.
- Collaborate with statisticians in the group and with scientists throughout the Lab and beyond (visiting faculty).
- Good mentorship.

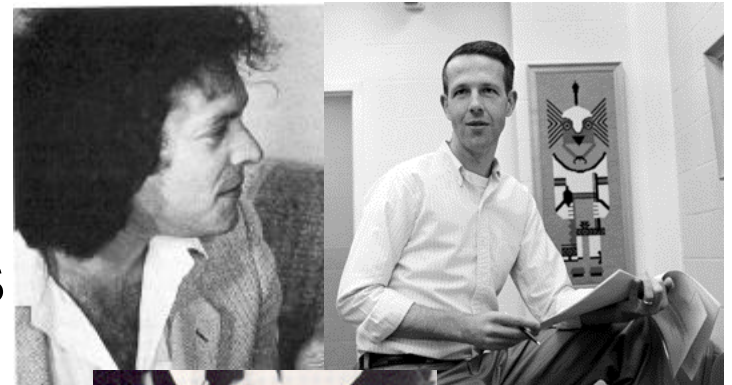




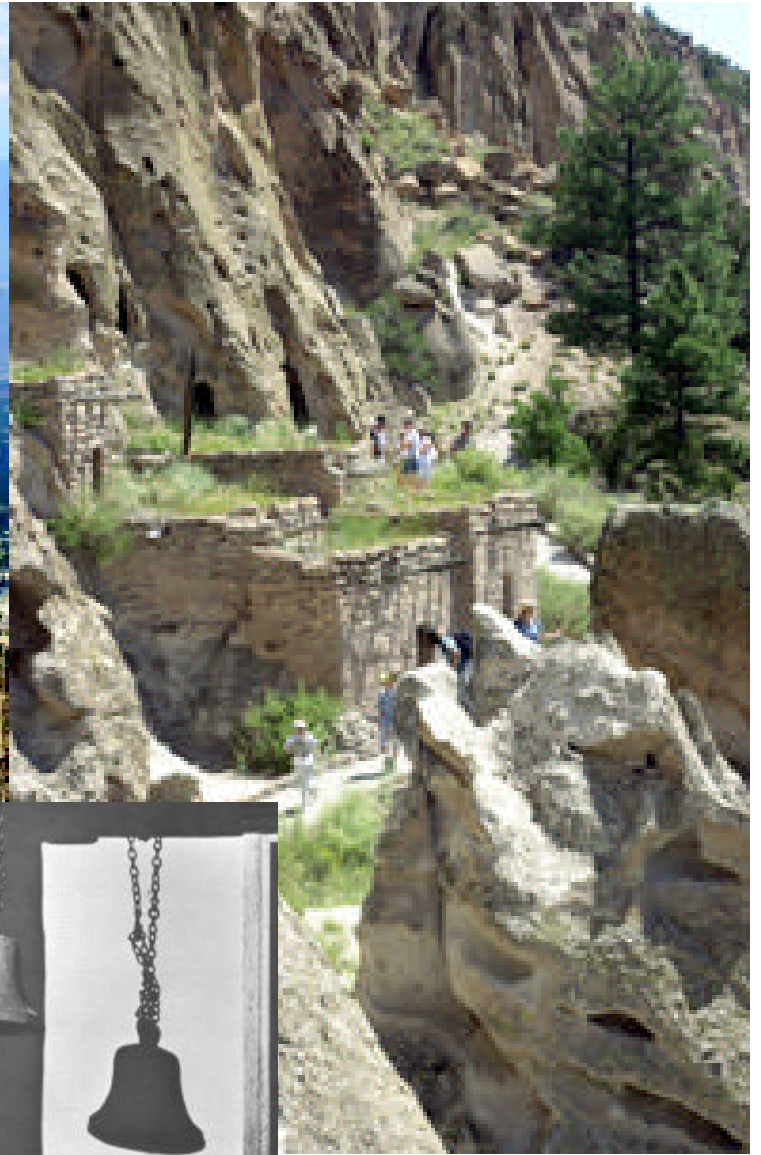


# We work with:

- Physicists
- Engineers
- Computer scientists (simulation modelers)
- Chemists
- Biologists
- Environmental scientists
- Applied mathematicians
- Each other!







Where  
we  
work!



# What We Do Well

---

- Reliability
- Information Integration Technology
- Theory and Methods for Computer Model Evaluation
- Statistical Population Bounding
- Monte Carlo Methods
- Computational Statistics
- Biological Sciences Applications





# Reliability

- Not boring!
- Historically: reliability of weapon systems
- Particularly important and challenging when some classes of tests are prohibited
- Now: reliability of manufacturing processes and design processes
- Reliability of world's fastest computer
- Competing Risk Models
- One-of-a-Kind Questions
- Hybrid Design





# Biological Sciences Applications

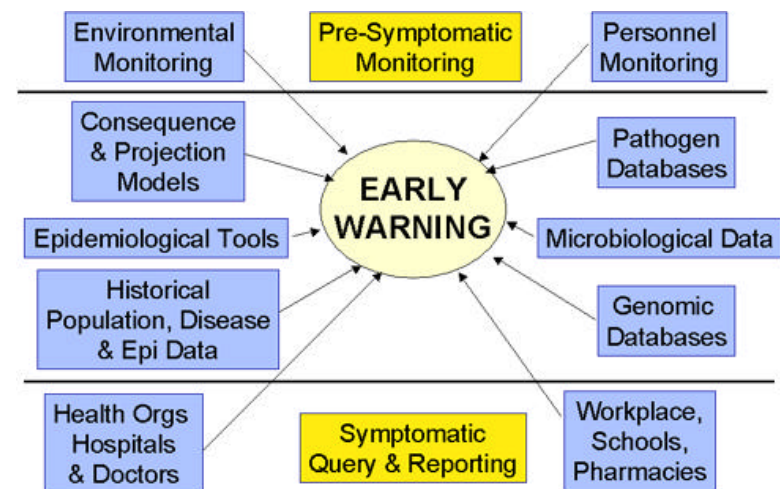
## Research Areas:

- Large-Scale Epidemiological Simulation (EpiSims)
- Genetic Data Analysis (AFLP Data)
- (National) Bio-Defense Initiative
- Ecological and Environmental Statistics

## Statistical Challenges:

- Information Integration
- Computer Model Evaluation
- Uncertainty Quantification
- Variance Estimation
- Spatio-Temporal Data

### Bio-Early Warning System:







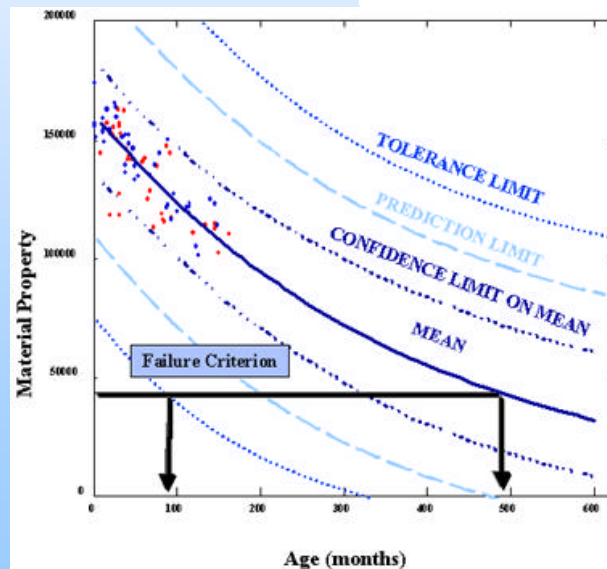
# Statistical Population Bounding

Tolerance limits provide bounds at which no more than a specified % of the population will fail.

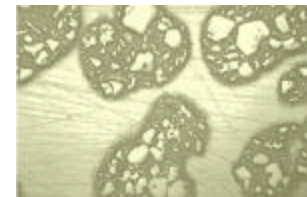
Confidence limits bound the mean function with a specified level of confidence.

Prediction limits bound individual predicted points.

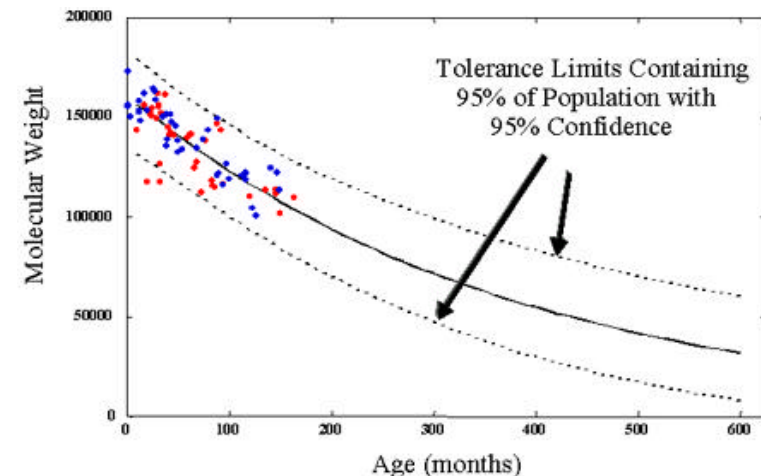
Tolerance bounds contain a specified proportion of a population with a desired confidence.



## Kinetics Model for Degradation of Estane Molecular Weight



Advanced Statistical Techniques Such As Computational Resampling Allow Population Bounding of the Predictions





# Information Integration Technology

Develop a “**standard**” framework of processes, methods and tools useful for evolving R&D to support decision making under uncertainty.





# Computational Statistics

---

- Existing tools are insufficient for laboratory methodology development
  - ◆ High dimensional spaces
  - ◆ Incorporation of importance sampling
  - ◆ Complex Bayesian methods
  - ◆ Convergence diagnostics and mixing improvements
- Need computational environments for rapid prototyping of new methods, particularly MCMC based methods
- Extensible object-oriented system (YADAS)





# Monte Carlo Methods

- Stochastic simulation for many processes of interest involves modeling them as Markov chains with suitably-defined continuous state spaces.
- Examples:
  - ◆ Simulating physical processes and studying statistical physics:
    - ★ Movement of pollutants, neutrons, or agents
  - ◆ Rare event simulation
  - ◆ Simulating from distributions with widely separated peaks
  - ◆ Metropolis algorithm with “wrong” equilibrium distribution for speed and adjusting the calculations afterwards: importance sampling

